

# VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS

## PCT


### INTERNATIONALER VORLÄUFIGER BERICHT ÜBER DIE PATENTIERBARKEIT

(Kapitel II des Vertrags über die internationale Zusammenarbeit auf dem Gebiet des Patentwesens)

REC'D 04 JAN 2006

WIPO

PCT

Aktenzeichen des Anmelders oder Anwalts 155394.2/Le	<b>WEITERES VORGEHEN</b> siehe Formblatt PCT/IPEA/416	
Internationales Aktenzeichen PCT/EP2004/053384	Internationales Anmeldedatum (Tag/Monat/Jahr) 09.12.2004	Prioritätsdatum (Tag/Monat/Jahr) 09.12.2003
Internationale Patentklassifikation (IPK) oder nationale Klassifikation und IPK G06F17/30		
Anmelder SWISS REINSURANCE COMPANY		
<p>1. Bei diesem Bericht handelt es sich um den internationalen vorläufigen Prüfungsbericht, der von der mit der internationalen vorläufigen Prüfung beauftragten Behörde nach Artikel 35 erstellt wurde und dem Anmelder gemäß Artikel 36 übermittelt wird.</p> <p>2. Dieser BERICHT umfaßt insgesamt 8 Blätter einschließlich dieses Deckblatts.</p> <p>3. Außerdem liegen dem Bericht ANLAGEN bei; diese umfassen</p> <p>a. <input checked="" type="checkbox"/> (an den Anmelder und das Internationale Büro gesandt) insgesamt 10 Blätter; dabei handelt es sich um</p> <p><input checked="" type="checkbox"/> Blätter mit der Beschreibung, Ansprüchen und/oder Zeichnungen, die geändert wurden und diesem Bericht zugrunde liegen, und/oder Blätter mit Berichtigungen, denen die Behörde zugestimmt hat (siehe Regel 70.16 und Abschnitt 607 der Verwaltungsvorschriften).</p> <p><input type="checkbox"/> Blätter, die frühere Blätter ersetzen, die aber aus den in Feld Nr. 1, Punkt 4 und im Zusatzfeld angegebenen Gründen nach Auffassung der Behörde eine Änderung enthalten, die über den Offenbarungsgehalt der internationalen Anmeldung in der ursprünglich eingereichten Fassung hinausgeht.</p> <p>b. <input type="checkbox"/> (nur an das Internationale Büro gesandt) insgesamt (bitte Art und Anzahl der/des elektronischen Datenträger(s) angeben), der/die ein Sequenzprotokoll und/oder die dazugehörigen Tabellen enthält/enthalten, nur in computerlesbarer Form, wie im Zusatzfeld betreffend das Sequenzprotokoll angegeben (siehe Abschnitt 802 der Verwaltungsvorschriften).</p>		
<p>4. Dieser Bericht enthält Angaben zu folgenden Punkten:</p> <p><input checked="" type="checkbox"/> Feld Nr. I Grundlage des Bescheids</p> <p><input type="checkbox"/> Feld Nr. II Priorität</p> <p><input type="checkbox"/> Feld Nr. III Keine Erstellung eines Gutachtens über Neuheit, erfinderische Tätigkeit und gewerbliche Anwendbarkeit</p> <p><input type="checkbox"/> Feld Nr. IV Mangelnde Einheitlichkeit der Erfindung</p> <p><input checked="" type="checkbox"/> Feld Nr. V Begründete Feststellung nach Artikel 35(2) hinsichtlich der Neuheit, der erfinderischen Tätigkeit und der gewerblichen Anwendbarkeit; Unterlagen und Erklärungen zur Stützung dieser Feststellung</p> <p><input type="checkbox"/> Feld Nr. VI Bestimmte angeführte Unterlagen</p> <p><input type="checkbox"/> Feld Nr. VII Bestimmte Mängel der internationalen Anmeldung</p> <p><input type="checkbox"/> Feld Nr. VIII Bestimmte Bemerkungen zur internationalen Anmeldung</p>		
Datum der Einreichung des Antrags  14.07.2005	Datum der Fertigstellung dieses Berichts  30.12.2005	
Name und Postanschrift der mit der internationalen Prüfung beauftragten Behörde   Europäisches Patentamt - Gitschiner Str. 103 D-10958 Berlin Tel. +49 30 25901 - 0 Fax: +49 30 25901 - 840	Bevollmächtigter Bediensteter  Polzer, A  Tel. +49 30 25901-427	



---

**Feld Nr. I Grundlage des Berichts**

---

1. Hinsichtlich der **Sprache** beruht der Bericht auf der internationalen Anmeldung in der Sprache, in der sie eingereicht wurde, sofern unter diesem Punkt nichts anderes angegeben ist.
- ☐ Der Bericht beruht auf einer Übersetzung aus der Originalsprache in die folgende Sprache, bei der es sich um die Sprache der Übersetzung handelt, die für folgenden Zweck eingereicht worden ist:
- ☐ internationale Recherche (nach Regeln 12.3 und 23.1 b))
  - ☐ Veröffentlichung der internationalen Anmeldung (nach Regel 12.4)
  - ☐ internationale vorläufige Prüfung (nach Regeln 55.2 und/oder 55.3)
2. Hinsichtlich der **Bestandteile**\* der internationalen Anmeldung beruht der Bericht auf *(Ersatzblätter, die dem Anmeldeamt auf eine Aufforderung nach Artikel 14 hin vorgelegt wurden, gelten im Rahmen dieses Berichts als "ursprünglich eingereicht" und sind ihm nicht beigelegt)*:

**Beschreibung, Seiten**

2, 4, 10, 13-21	in der ursprünglich eingereichten Fassung
1, 3, 5-9, 11, 12	eingegangen am 09.07.2005 mit Schreiben vom 06.07.2005

**Ansprüche, Nr.**

1-12, 15-24	in der ursprünglich eingereichten Fassung
13, 14	eingegangen am 09.07.2005 mit Schreiben vom 06.07.2005

**Ansprüche, Seiten**

22-24, 26-28	in der ursprünglich eingereichten Fassung
25	eingegangen am 09.07.2005 mit Schreiben vom 06.07.2005

**Zeichnungen, Blätter**

1/13-13/13	in der ursprünglich eingereichten Fassung
------------	-------------------------------------------

☐ einem Sequenzprotokoll und/oder etwaigen dazugehörigen Tabellen - siehe Zusatzfeld betreffend das Sequenzprotokoll

3. ☐ Aufgrund der Änderungen sind folgende Unterlagen fortgefallen:
- ☐ Beschreibung: Seite
  - ☐ Ansprüche: Nr.
  - ☐ Zeichnungen: Blatt/Abb.
  - ☐ Sequenzprotokoll (*genaue Angaben*):
  - ☐ etwaige zum Sequenzprotokoll gehörende Tabellen (*genaue Angaben*):

4. ☐ Dieser Bericht ist ohne Berücksichtigung (von einigen) der diesem Bericht beigelegten und nachstehend aufgelisteten Änderungen erstellt worden, da diese aus den im Zusatzfeld angegebenen Gründen nach Auffassung der Behörde über den Offenbarungsgehalt in der ursprünglich eingereichten Fassung hinausgehen (Regel 70.2 c)).

- ☐ Beschreibung: Seite
- ☐ Ansprüche: Nr.
- ☐ Zeichnungen: Blatt/Abb.
- ☐ Sequenzprotokoll (*genaue Angaben*):
- ☐ etwaige zum Sequenzprotokoll gehörende Tabellen (*genaue Angaben*):

*\* Wenn Punkt 4 zutrifft, können einige oder alle dieser Blätter mit der Bemerkung "ersetzt" versehen werden.*

---

**Feld Nr. V Begründete Feststellung nach Artikel 35 (2) hinsichtlich der Neuheit, der erfinderischen Tätigkeit und der gewerblichen Anwendbarkeit; Unterlagen und Erklärungen zur Stützung dieser Feststellung**

---

1. Feststellung
- |                                |                                           |
|--------------------------------|-------------------------------------------|
| Neuheit (N)                    | Ja: Ansprüche 3,7-11,14,18-22             |
|                                | Nein: Ansprüche 1,2,4-6,12,13,15-17,23,24 |
| Erfinderische Tätigkeit (IS)   | Ja: Ansprüche                             |
|                                | Nein: Ansprüche 1-24                      |
| Gewerbliche Anwendbarkeit (IA) | Ja: Ansprüche 1-24                        |
|                                | Nein: Ansprüche:                          |

2. Unterlagen und Erklärungen (Regel 70.7):

**siehe Beiblatt**

**Zu Punkt V**

Es wird auf folgende Dokumente verwiesen:

D1: US 2003/195872 A1 (SENN PAUL) 16. Oktober 2003 (2003-10-16)

D2: DAS S R ET AL: "Yahoo! for Amazon: Sentiment Parsing from Small Talk on the Web" EFA 2001 BARCELONA MEETINGS, [Online] 5. August 2001 (2001-08-05), Seiten 1-45, XP002324570 Gefunden im Internet:  
URL:[http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=276189](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=276189)> [gefunden am 2005-04-13]

1. Die vorliegende Anmeldung erfüllt nicht die Erfordernisse des Artikels 33(1) PCT, weil der Gegenstand der Ansprüche 1-24 nicht neu im Sinne von Artikel 33(2) PCT ist, oder nicht auf einer erfinderischen Tätigkeit im Sinne von Artikel 33(3) beruht.
- 1.1. Das Dokument D1 offenbart (die Verweise in Klammern beziehen sich auf dieses Dokument):

Verfahren zur Aggregation und Analyse von dezentralisiert gespeicherten Multimediadaten, wobei in einem Datenspeicher ein oder mehrere verknüpfbare Suchbegriffe abgespeichert werden, eine Recheneinheit über ein Netzwerk auf mit Quelldatenbanken verbundene Netzwerknodes zugreift und Daten der Quelldatenbanken basierend auf den Suchbegriffen selektiert werden (übliches Verfahren bei Internet Suchmaschinen, siehe z.B. D1 Absätze [0017], [0018], Abbildung 1), dadurch gekennzeichnet,

- a. dass in einem Datenspeicher mindestens ein Wertungsparameter einem Suchbegriff und/oder einer Verknüpfung von Suchbegriffen zugeordnet abgespeichert wird (Absatz [0007] Zeilen 4-5: Suchbegriff "Microsoft" verknüpft mit "at least one word with positive emotional connotations"; Absatz [0088]: "combination" von "words retrieved in step 2" und "user-chosen search keys". In "step 2"/Absatz [0087] werden Wertungsparameter mit bestimmter emotionaler Konnotation ausgewählt, vgl. Absatz [0072]),
- b. dass im Datenspeicher mindestens eine der Quelldatenbanken einem Suchbegriff und/oder einer Verknüpfung von Suchbegriffen zugeordnet

- abgespeichert wird (in Absatz [0086] wird ein benutzerdefinierter Parameter "where to perform the search (internet, intranet, etc.)" in Zusammenhang mit der Eingabe der Suchbegriffe erwähnt, was einer Einschränkung der Suche auf bestimmte Webserver entspricht, siehe Absatz [0088] erste zwei Zeilen. Laut Absatz [0019] viertletzte Zeile entspricht die Suche des Verfahrens von D1 herkömmlichen Suchmaschinen wie AltaVista, bei denen die Möglichkeit der Einschränkung der Suche auf bestimmte Webserver natürlich ein seit Jahren übliches Merkmal ist. Zur Einschränkung der Suche auf bestimmte Diskussionsforen im Internet siehe auch D2, Abschnitt 5 zweiter Absatz),
- c. dass mittels eines Filtermoduls ("Content Retrieval Object", Absätze [0085] - [0089]) der Recheneinheit auf die Quelldatenbanken der Netzwerknodes zugegriffen wird und für jeden Wertungsparameter in Verbindung mit den zugeordneten Suchbegriffen (Absatz [0079]) und den zugeordneten Quelldatenbanken (Absatz [0081]) und/oder einer zeitlichen Wertung der Dokumente (Absatz [0080]. Das fakultative Merkmal "Zeitliche Wertung" ist extrem ungenau und wird im Zusammenhang mit der Generierung der Wertungsliste auch in der Beschreibung nicht näher ausgeführt. "Dokumente" wurden in diesem Anspruch bisher nicht erwähnt) eine Wertungsliste mit gefundenen Datensätzen erzeugt wird (Absätze [0088], [0089]) und
- d. dass mittels eines Parametrisierungsmoduls ("Content Analysis Engine", Absätze [0109] und folgende) basierend auf der Wertungsliste für den jeweiligen Wertungsparameter eine variable Stimmungsgrösse mindestens teilweise dynamisch generiert wird (Absatz [0226], "composite rating"), welche variable Stimmungsgrösse zeitlichen Stimmungsschwankungen von Benutzern des Netzwerkes entsprechen (dies kann nicht als technisches Merkmal angesehen werden, sondern ist ein durch die Erfindung zu erzielender Effekt, siehe Abschnitt 2.1. Das in D1 beschriebene Verfahren verfolgt jedoch einen ähnlichen Zweck, insbesondere werden auch zeitliche Stimmungsschwankungen verfolgt, siehe Absatz [0226] letzte sechs Zeilen).

Der Gegenstand des **Anspruches 1** kann daher nicht als neu betrachtet werden.

- 1.2. Die während der Internationalen Vorläufigen Prüfung vorgebrachten Argumente vermögen aus den folgenden Gründen nicht die Neuheitsargumentation unter 1.1 zu



widerlegen:

Es wurde argumentiert dass nach Anspruch 1 nur eine globale Stimmungsgrosse generiert werde, wohingegen in D1 lediglich ein "composite rating" für jeden gefundenen website erstellt werde. Dies ist jedoch nicht zutreffend. Nach D1 Absätze [0028] und [0035] werden die "composite ratings" für "ratable units" erstellt, wobei die "ratable units" rekursiv definiert werden können, z.B. kann ein Satz aus mehreren grammatikalischen Einheiten und ein Absatz aus mehreren Sätzen bestehen, siehe D1 Absatz [0117] Zeilen 14-22. Insbesondere besteht ein "search result object" als "ratable unit" aus den in einer Internetsuche gefundenen Webseiten, siehe D1 Absatz [0117] Zeilen 25-26 sowie Absätze [0088] und [0089]. In D1 Absatz [0117] Zeilen 26-34 sowie Absätzen [0197] und folgenden wird allgemein erläutert wie die Bewertung einer "ratable unit" rekursiv aus den Bewertungen ihrer Bestandteile gebildet wird, z.B. durch Durchschnitts- oder Medianbildung. Nach dem oben gesagten betrifft dies implizit klar auch die Bildung einer Bewertung für das "search result object" aus den Bewertungen der in der Internetsuche gefundenen Webseiten. In D1 Absatz [0226] wird schließlich als Ergebnis des Algorithmus ein "SearchResult object ... containing the Composite Rating for the search" erwähnt. Daraus und aus der Darstellung in Absatz [0226] Zeilen 14-23 des Vergleichs zwischen den "composite ratings" für Internetsuchen mit den Suchargumenten "Sun" und "Microsoft" geht klar hervor dass das System von D1 für jede Internetsuche einen einzelnen Zahlenwert als "composite rating" generiert, der der "Stimmungsgrosse" von Anspruch 1 entspricht.

Es wurde weiterhin argumentiert, aus dem "composite rating" von D1 gehe keine Information über die Stimmungsschwankung von Netzwerkbenutzern hervor und die Erfindung nach Anspruch 1 ermögliche im Gegensatz zu D1 die zeitliche Verfolgung von Stimmungsschwankungen. Dies ist im Anspruch 1 nicht als technisches Merkmal formuliert, siehe die Argumentation zu Schritt c von Anspruch 1 unter 1.1. Insbesondere erwähnt jedoch D1 in Absatz [0226] Zeilen 14-23 explizit eine zeitliche Verfolgung von Stimmungsschwankungen auf dem Netzwerk durch wiederholte Suche nach zwei Suchbegriffen und Vergleich des damit jeweils verbundenen "emotional tone" in einem "contrast index". In dem Beispiel von D1 ergibt sich als Ergebnis "Microsoft is gaining in popularity", was sich klar auf den Vergleich der

globalen Einstellung zur Firma Microsoft auf dem Internet zu zwei unterschiedlichen Zeitpunkten bezieht.

Weitere Neuheitsargumente des Anmelders zur Bildung des "composite ratings" von D1 aus den Bewertungen von hierarchisch zusammengesetzten Teileinheiten beziehen sich offensichtlich nicht auf Merkmale des Anspruches 1, der in Schritt d keine Details zur Bildung der Stimmungsgrösse enthält; sie können daher die Neuheit des Anspruches ebenfalls nicht stützen.

Da Anspruch 1 somit nicht neu ist erübrigt sich ein Eingehen auf Argumente zur erfinderischen Tätigkeit.

- 1.3. Die abhängigen Ansprüche 2-11 enthalten aus den folgenden Gründen keine Merkmale, die in Kombination mit den Merkmalen irgendeines Anspruchs, auf den sie sich beziehen, die Erfordernisse des PCT in bezug auf Neuheit bzw. erfinderische Tätigkeit erfüllen:

**Anspruch 2:** Abspeichern und Anzeigen von Ergebnislisten sind übliche Merkmale von Internet Suchmaschinen und sind in D1 impliziert, siehe z.B. D1 Absätze [0089] und [0226]. Der Anspruch ist daher nicht neu.

**Anspruch 3:** Zum periodischen Suchen zum Verfolgen von Veränderungen in den Suchergebnissen siehe D1 Absatz [0226] zweite Hälfte. Nötigenfalls würde der Fachmann auch D1 mit Dokument D2 kombinieren in dem die periodische Suche zum Verfolgen von globalen Stimmungsschwankungen auf dem Internet ausführlich diskutiert wird, siehe D2 Seite 26 Zeile 9 - Seite 27 Zeile 3 und Abbildungen 2-7. Ein Alarmmechanismus zur Benachrichtigung des Benutzers wenn Suchergebnisse bestimmte Bedingungen erfüllen ist im Fachgebiet weithin bekannt und wäre für den Fachmann naheliegend. Der Anspruch ist daher nicht erfinderisch.

**Anspruch 4:** Der Anspruch ist nicht neu, da er durch das "Information Dimension Dictionary" von D1 antizipiert wird, s. D1 Absätze [0070], [0072] und [0087].

**Anspruch 5:** Die dynamische Erzeugung des "Information Dimension Dictionary"

wird in D1 Absätze [0093] - [0098] beschrieben, der Anspruch ist daher nicht neu.  
Siehe auch Abschnitt 2.2.

**Anspruch 6:** Die Verwendung der erwähnten Standards ist bei Internet Suchmaschinen gang und gäbe, zumindest HTML wird klar in D1 impliziert. Der Anspruch ist daher nicht neu.

**Ansprüche 7-9:** Die Ansprüche sind nicht erfinderisch, da die Anpassung der Darstellung von Suchergebnissen an Benutzerprofile oder an Eigenschaften von vom Benutzer verwendeten Endgeräten sowie das automatische Anpassen von Benutzerprofilen durch Beobachten des Benutzerverhaltens im Gebiet der Internet Suchmaschinen extrem verbreitet sind und für den Fachmann naheliegend wären.

**Ansprüche 10, 11:** Das durch die in den Ansprüchen enthaltenen Merkmale gelöste Problem ist die Ermöglichung der Evaluation von historischen Werten sowie die Vorhersage von zukünftigen Werten der Stimmungsgrösse. Der Fachmann würde bereits in dem Verfahren von D1 auf dieses Problem stossen, da in D1 Absatz [0226] letzte sechs Zeilen bereits das Verfolgen von zeitlichen Schwankungen der Stimmungsgrösse thematisiert wird. Nötigenfalls würde der Fachmann auch D1 mit Dokument D2 kombinieren in dem die periodische Suche zum Verfolgen von globalen Stimmungsschwankungen auf dem Internet ausführlich diskutiert wird, siehe D2 Seite 26 Zeile 9 - Seite 27 Zeile 3 und Abbildungen 2-7. Das Abspeichern historischer Werte sowie die Verwendung eines Extrapolationsmoduls zur Vorhersage zukünftiger Werte wären für den Fachmann naheliegende Merkmale zur Lösung des genannten Problems, die Ansprüche sind daher nicht erfinderisch.

- 1.4. Die **Vorrichtungsansprüche 12-22** sowie **Computerprogrammansprüche 23 und 24** enthalten technische Merkmale die den Merkmalen der Verfahrensansprüche 1-11 entsprechen. Die Argumentation unter 1.1 and 1.3 gilt daher für diese Ansprüche *mutatis mutandis*.



## **System und Verfahren zur Aggregation und Überwachung von dezentralisiert gespeicherten Multimediadaten**

Die Erfindung betrifft ein System und ein Verfahren zur Aggregation und Analyse von dezentralisiert gespeicherten Multimediadaten, wobei in einem  
5 Datenspeicher ein oder mehrere verknüpfbare Suchbegriffe abgespeichert werden, eine Recheneinheit über ein Netzwerk auf mit Quelldatenbanken verbundene Netzwerknodes zugreift und Daten der Quelldatenbanken basierend auf den Suchbegriffen selektiert werden. Die Erfindung betrifft insbesondere ein System und Verfahren zur Realtime-Analyse solcher  
10 dezentralisiert gespeicherter Multimediadaten.

Das Internet bzw. das weltweite Backbone-Netz ist heute zweifelsohne eine der wichtigsten Quellen zur Informationsbeschaffung in Industrie, Wissenschaft und Technik und gehört wahrscheinlich zu den wichtigsten technischen Errungenschaften des ausgehenden 20. Jahrhunderts.  
15 Es ist eine Tatsache, dass heute über das Internet auf gigantische Datenmengen zugegriffen werden kann, in einem Ausmass, wie es bis vor 10 Jahren noch kaum vorstellbar war. Trotz all den daraus entstehenden Vorteilen ergibt sich damit jedoch auch die Schwierigkeit, wirklich relevante Daten in dieser riesigen Datenmenge aufzufinden. Search-Engines, wie z.B. die  
20 bekannten Internet-Search-Engines mit z.B. dem bekannten Altavista-Engine als wortbasierende Suchmaschine oder z.B. der Yahoo-Engine als topicbasierende Suchmaschine, machen die Vielzahl der dezentralisierten Datenquellen für den Benutzer erst nutzbar, da ohne solche Hilfsmittel die Aussicht, dass möglichst viele der relevanten Daten wirklich gefunden werden,  
25 drastisch sinkt. Es kann gesagt werden, dass das Internet ohne Search-Engines wie ein Kraftfahrzeug ohne Motor ist. Dies zeigt sich insbesondere in der statistischen Tatsache, dass die Benutzer des Internets mehr Online-Zeit bei Search-Engines verbringen, als irgendwo sonst. Trotz allen Fortschritten auf diesem Gebiet gibt die im Stand der Technik verfügbare Search-Engine-  
30 Technologie dem Benutzer jedoch häufig keine wirklich zufrieden stellenden Antworten. Als Beispiel sei angenommen, ein Benutzer möchte Informationen, z.B. zum Auto Model-Typ Fiat Uno finden, z.B. in Zusammenhang mit einer Haftungsklage einer Produkthaftung in Bezug auf ein fehlerhaftes Design mit

Funktionsweise von Robots/Crawlers. Search-Robots oder Crawlers zeichnen sich durch einen Prozess aus (d.h. den Crawler), welcher sich durch das Netzwerk 70, hier das Internet 701-704, von Netzwerk-Node 73 zu Netzwerk-Node 73 bzw. von Web-Site 73 zu Web-Site 73 bewegt (Pfeil 71) und dabei  
5 den Inhalt jedes Web-Dokumentes, welches er findet, an seinen Host-Rechner 72 zurückschickt. Der Host-Rechner 72 indexiert die durch den Crawler geschickten Web-Dokumente 722 und speichert die Information in einer Datenbank 721 ab. Jede Suchanfrage (Request) durch einen Benutzer greift auf die Informationen der Datenbank 721 zu. Die Crawlers des Standes der  
10 Technik betrachten normalerweise jede Information als relevant, weshalb alle irgendwo gefundenen Web-Dokumente durch den Host-Rechner 72 indexiert werden. Beispiele solcher Robots/Crawlers sind u.a. Google<sup>TM</sup>, Altavista<sup>TM</sup> und Hotbot<sup>TM</sup>. Figur 2 illustriert die sog. Metacrawlers. Metacrawlers unterscheiden sich von den Robots/Crawlers durch die Möglichkeit, mittels einer einzigen  
15 Sucheinrichtung 82 zu suchen, wobei die Antwort zusätzlich durch eine Vielzahl von weiteren Systemen 77 des Netzes 75 erzeugt wird. Der Metacrawler dient somit als ein Front-End zu einer Vielzahl von weiteren Systemen 77. Die Antwort auf einen Suchrequest von einem Metacrawler wird typischerweise durch die Anzahl seiner weiteren Systeme 77 begrenzt. Beispiele von  
20 Metacrawlers sind u.a. MetaCrawler<sup>TM</sup>, LawCrawler<sup>TM</sup> und LawRunner<sup>TM</sup>.

Kataloge mit oder ohne Suchmöglichkeiten zeichnen sich durch eine spezielle Auswahl von Links aus, welche von Hand strukturiert und/oder organisiert und in einer entsprechenden Datenbank abgespeichert werden. Im Fall eines Kataloges mit Suchmöglichkeiten wird bei einem Suchrequest die  
25 manuell gespeicherte Information durch das System nach den gewünschten Suchtermen abgesucht. Im Fall eines Kataloges ohne Suchmöglichkeiten muss der Benutzer die gewünschte Information selbst aus der Liste der gespeicherten Links suchen, indem er z.B. manuell durch die Liste klickt oder scrollt. Im letzteren Fall entscheidet der Benutzer selbst, welche Information  
30 aus der Liste ihm relevant und welche ihm weniger relevant erscheint. Kataloge sind natürlicherweise durch das Leistungsvolumen und die Prioritäten des/der Editor(en) begrenzt. Beispiele solcher Kataloge sind u.a. Yahoo!<sup>TM</sup> und FindLaw<sup>TM</sup>. Kataloge fallen unter die Kategorie der Portale und/oder Vortale. Portale und bis zu einem gewissen Mass z.B. auch proprietäre Datenbanken

Dokument findet, speichert es sie in der Datenbank des Crawlers, so dass der Crawler die entsprechenden Dokumente später ebenfalls in den Datenspeicher laden kann. Wie der Inhalt des Dokumentes indexiert wird, hängt von der jeweiligen Suchmaschine ab. Die indexierte Information kann z.B. in einer Hash-Table oder einem anderen geeigneten Tool zur späteren Verwendung gespeichert werden. Ein Benutzer kann nun ein Suchrequest über das Front-End eingeben und der Search-Engine sucht die entsprechenden indexierten Seiten. Das Verfahren basiert auf den Prinzip "Alles ist relevant", was bedeutet, dass der Crawler jedes Web-Dokument holen und speichern wird, dass  
irgendwie zugreifbar ist. Komplexe, contentorientierte Abfragen sind mit den heutigen Suchmaschinen nicht durchführbar, ohne dass sie entweder relevante Dokumente ausschliessen oder eine Flut von für die Abfrage irrelevanten Dokumenten mitangeben. Gerade bei Suchabfragen, bei welchen Themen basierend auf themenfremden, nicht scharf fassbaren Parametern indexiert werden sollen, ergeben die Search-Engines kaum je auch nur annähernd befriedigende Antworten. Wie erwähnt kann als ein Beispiel dazu das für die Industrie eminent wichtige Problem angeführt werden, dass zu einem konkreten Thema allgemeine Stimmungstendenzen, Meinungstendenzen oder Stimmungsschwankungen der Benutzer des Netzwerkes erfasst werden sollen. Dies ist basierend auf den heutigen Suchmaschinen nicht durchführbar. Ebenfalls ist es mit den Search-Engines des Standes der Technik bis anhin in keiner Weise möglich, Stimmungen und Stimmungsschwankungen der Netzwerkbenutzer zu einem Thema frühzeitig zu erkennen und die entsprechenden Dokumente anzugeben.

Die US-Patentanmeldung US2003/0195872 offenbart ein System, welches dazu benutzt werden kann, Suchbegriffe mit emotionellen Wertungsbegriffen zu verbinden und eine Suche im Internet und/oder Intranet basierend auf dieser Zuordnung von Suchbegriffen und emotionellen Wertungsbegriffen durchzuführen. Das System erlaubt jedoch kein gezieltes Screening von Datenbanken. Insbesondere sind keine zeitlichen Aussagen mittels des Systems möglich. Dies verhindert bzw. verunmöglicht eine objektive Beurteilung von Tendenzen oder zu erwartenden Ereignissen. Das System erlaubt lediglich ein statisches Listing von in den verfügbaren Datenbanken gespeicherten Dokumenten. Damit müssen alle relevanten Dokumente in



diesem System doch mehr oder weniger vollständig nach dem Listing gelesen und interpretiert werden, was eine Automatisierung im Sinne z.B. eines dynamischen Warnsystems verunmöglicht.

Es ist eine Aufgabe dieser Erfindung, ein neues System und ein  
5 Verfahren zur Aggregation und Analyse von dezentralisiert gespeicherten  
Multimediadaten vorzuschlagen, welche die oben genannten Nachteile des  
Standes der Technik nicht aufweisen. Insbesondere soll ein automatisiertes,  
einfaches und rationelles System und Verfahren vorgeschlagen werden, um  
komplexe, contentorientierte Abfragen durchzuführen. Bei der Abfrage sollen  
10 insbesondere themenfremden und/oder nicht scharf fassbaren Parametern, wie  
z.B. Stimmungen oder Stimmungsschwankungen der Netzbenutzer, als  
Filterparameter möglich sein. Umgekehrt sollen mit dem erfindungsgemässen  
Verfahren und System ebenfalls möglich sein, Stimmungen und  
Stimmungsschwankungen der Netzwerkbenutzer zu einem Thema frühzeitig zu  
15 erkennen und die entsprechenden Dokumente anzugeben.

Gemäss der vorliegenden Erfindung wird dieses Ziel insbesondere  
durch die Elemente der unabhängigen Ansprüche erreicht. Weitere vorteilhafte  
Ausführungsformen gehen ausserdem aus den abhängigen Ansprüchen und  
der Beschreibung hervor.

20 Insbesondere werden diese Ziele durch die Erfindung dadurch er-  
reicht, dass zur Aggregation und Überwachung und/oder Analyse von  
dezentralisiert gespeicherten Multimediadaten in einem Datenspeicher ein oder  
mehrere verknüpfbare Suchbegriffe abgespeichert werden, eine Recheneinheit  
über ein Netzwerk auf mit Quelldatenbanken verbundene Netzwerknodes  
25 zugreift und Daten der Quelldatenbanken basierend auf den Suchbegriffen  
selektiert werden, dass in einem Datenspeicher mindestens ein  
Wertungsparameter einem Suchbegriff und/oder einer Verknüpfung von  
Suchbegriffen zugeordnet abgespeichert wird, dass im Datenspeicher  
mindestens eine der Quelldatenbanken einem Suchbegriff und/oder einer  
30 Verknüpfung von Suchbegriffen zugeordnet abgespeichert wird, dass mittels  
eines Filtermoduls der Recheneinheit auf die Quelldatenbanken der  
Netzwerknodes zugegriffen wird und für jeden Wertungsparameter in

Verbindung mit den zugeordneten Suchbegriffen und den zugeordneten Quelldatenbanken und/oder einer zeitlichen Wertung der Dokumente eine Wertungsliste mit gefundenen Datensätzen erzeugt wird und dass mittels eines Parametrisierungsmoduls basierend auf der Wertungsliste für den jeweiligen Wertungsparameter eine variable Stimmungsgrösse mindestens teilweise dynamisch generiert wird, welche variable Stimmungsgrösse zeitliche, positive und/oder negative Stimmungsschwankungen von Benutzern des Netzwerkes entsprechen. Die Recheneinheit kann z.B. zur Generierung der variablen Stimmungsgrößen und/oder der Daten des Contentmoduls ein HTML- (Hyper Text Markup Language) und/oder HDML- (Handheld Device Markup Language) und/oder WML- (Wireless Markup Language) und/oder VRML- (Virtual Reality Modeling Language) und/oder ASP (Active Server Pages) -Modul umfassen. Diese Ausführungsvariante hat u.a. den Vorteil, dass das System auf einer vorgängig spezifisch definierbaren Gesamtheit an Quellen aus einem Netzwerk, insbesondere aus dem Internet (z.B. Web-Sites, Chat Rooms, E-mail Foren etc.) basiert, welche ebenfalls nach vorgängig definierbaren Suchkriterien gescannt werden. Das System ermöglicht also nicht nur die Generierung einer "Trefferliste" von im Internet gefundenen Web-Sites mit entsprechendem Inhalt, sondern das System ermöglicht das erwähnte Screening von vordefinierbaren Quellen und deren systematische und dadurch quantitativ relevante Auswertung entsprechend den gewünschten und definierten inhaltlichen Kriterien (z.B. welche Medikamente werden im Zusammenhang mit schweren Nebenwirkungen genannt - und die in welcher Häufigkeit). Dieses inhaltliche Screening kann in einer periodischen Abfolge (zeitlich) erfolgen, wobei sämtliche gefunden 'Treffer'-Inhalte jederzeit wieder verfügbar gemacht werden können und somit statistische Aussagen, gerade eben über die Zeit, möglich sind. Natürlich können die Dokumente auch anderweitig auf ihre zeitliche Zuordnung, z.B. basierend auf dem Abspeicherungsdatum erfasst werden. Das System erkennt also auch wann welcher Inhalt in den besagten Quellen abgelegt wurde. Dadurch dass auf diese Art eine quantitative Auswertung möglich wird kann das System die definierten Quellen selbständig 'monitoren' und ein überschreiten eines 'Treshold-Wertes' (quantitativ) entsprechend darstellen. Das System ermöglicht es, Suchkriterien so zu definieren, dass nach einem inhaltliche logischen Zusammenhang (einen Sinn ergebend) gesucht werden kann (nicht nur das



Stichwort zählt, sondern inhaltlicher Zusammenhang). Das System verbindet damit die Suchkriterien zu einem Inhalt, nach welchen dann gesucht wird.

In einer Ausführungsvariante werden ein oder mehrere der Wertungsparameter mittels einer lexikographischen Wertungsdatenbank generiert werden. Dasselbe kann für die Suchbegriffe realisiert sein. Diese Ausführungsvariante hat u.a. den Vorteil, dass Such- und Wertungsbegriffe benutzerspezifisch und/oder applikationsspezifisch definiert werden können. Als Ausführungsvariante kann die lexikographischen Wertungsdatenbank und/oder Suchbegriffdatenbank dynamisch basierend auf bereits durchgeführten Suchen/Analysen ergänzt und/oder verändert werden. Damit kann das System an veränderte Bedingungen und/oder Wortbildungen automatisiert angepasst werden, was so im Stand der Technik nicht möglich war.

In einer anderen Ausführungsvariante werden ein oder mehrere der Wertungsparameter dynamisch mittels der Recheneinheit während der Erzeugung der Wertungsliste generiert. Diese Ausführungsvariante hat u.a. die gleichen Vorteile wie die vorhergehenden Ausführungsvariante.

In einer weiteren Ausführungsvariante wird die Wertungsliste mit den gefundenen Datensätzen und/oder Verweisen auf die gefundenen Datensätze in einem Contentmodul der Recheneinheit für einen Benutzer zugreifbar abgespeichert. Diese Ausführungsvariante hat u.a. den Vorteil, dass das System z.B. als ein Warnsystem für den Benutzer eingesetzt werden kann, das ihn über bevorstehenden Tendenzen im Markt oder der Bevölkerung ( z.B. Class Actions etc.) informiert und/oder warnt.

In einer Ausführungsvariante werden periodisch die Stimmungsgrößen mittels der Recheneinheit überprüft und falls mindestens eine der Stimmungsgrößen ausserhalb einer festlegbaren Schwankungstoleranz oder bestimmbarer Erwartungswertes liegt, die entsprechende Wertungsliste mit den gefundenen Datensätze und/oder Verweisen auf gefundene Datensätze im Contentmodul der Recheneinheit für einen Benutzer zugreifbar abgespeichert und/oder aktualisiert. Diese

Ausführungsvariante hat u.a. den Vorteil, dass die Datenbanken auf zeitliche Veränderungen oder zu erwartende Ereignisse, z.B. mittels definierbarem Wahrscheinlichkeitsschwellwert, gezielt gescannt werden können und den Benutzer so z.B. rechtzeitig warnen können (z.B. Produktfehler, Produkthaftung etc.)

In einer wieder anderen Ausführungsvariante wird ein Benutzerprofil anhand von Benutzerinformationen erstellt, wobei basierend auf den im Contentmodul abgespeicherten gefundenen Datensätzen und/oder Verweisen auf gefundene Datensätze mittels eines Repackagingmoduls unter Berücksichtigung der Daten des Benutzerprofils benutzerspezifisch optimierte Daten erzeugt werden, welche benutzerspezifisch optimierten Daten dem Benutzer im Contentmodul der Recheneinheit abgespeichert zur Verfügung gestellt werden. Dem Benutzer können als Ausführungsvariante verschiedene Benutzerprofile für unterschiedliche Kommunikationsvorrichtungen des Benutzers zugeordnet abgespeichert werden. Weiter können z.B. auch Daten zum Benutzerverhalten von der Recheneinheit automatisch erfasst und dem Benutzerprofil zugeordnet abgespeichert werden. Diese Ausführungsvariante hat u.a. den Vorteil, dass unterschiedliche Accessmöglichkeiten des Benutzers benutzerspezifisch berücksichtigt werden können und das System so benutzerspezifisch optimiert werden kann.

In einer Ausführungsvariante werden mittels eines Historymoduls zu jeder berechneten variablen Stimmungsgrösse die Werte bis zu einem festlegbaren vergangenen Zeitpunkt abgespeichert. Diese Ausführungsvariante hat u.a. die gleichen Vorteile einer zeitlichen Kontrolle und Erfassung von Veränderungen innerhalb der abgespeicherten und zugreifbaren Dokumenten.

In einer weiteren Ausführungsvariante berechnet die Recheneinheit mittels eines Extrapolationsmoduls Erwartungswerte zu einer bestimmaren Stimmungsgrösse basierend auf den Daten des Historymoduls für einen bestimmaren zukünftigen Zeitpunkt und speichert sie in einem Datenspeicher der Recheneinheit ab. Diese Ausführungsvariante hat u.a. den Vorteil, dass zu erwartende Ereignisse automatisiert vorausgesagt werden können. Dies kann nicht nur bei Warnsystemen (z.B. gegen Class Actions bei Produkthaftung

ein oder mehrere verknüpfbare Suchbegriffe 310,311,312,313 abgespeichert.  
Eine Recheneinheit 10 greift über ein Netzwerk 50 auf mit Quelldatenbanken  
401,411,421,431 verbundene Netzwerknodes 40,41,42,43 zu und Daten der  
Quelldatenbanken 401,411,421,431 basierend auf den Suchbegriffen  
5 310,311,312,313 werden selektiert.

Figur 4 zeigt beispielhaft ein mögliches Ergebnis bei einem  
medizinischen und/oder pharmazeutischen Überwachungssystem basierend  
auf Medikamente in Abhängigkeit ihrer Trefferliste in den Dokumenten.

Figur 5 zeigt ebenfalls in einem solchen medizinischen und/oder  
10 pharmazeutischen Überwachungssystem beispielhaft ein mögliches Ergebnis  
z.B. eines Medikamentes in Zusammenhang mit auftretenden Krankheiten  
und/oder Todesursachen.

Figur 6 zeigt in der gleichen Ausführungsvariante von Figur 4 und 5  
das zeitlich erfasste Auftreten am Beispiel von Serzone in den Dokumenten der  
15 verfügbaren und/oder bestimmten Quelldatenbanken 401,411,421,431.

Figur 7 zeigt ein beispielhaftes Listing von Firmen (hier z.B.  
Anwaltskanzleiseiten etc.) in Abhängigkeit von einer Selektion von Wertungs-  
und/oder Suchbegriffen 310,311,312,313 (hier z.B. Industrienamen) und ihrer  
Trefferzahl in den Dokumenten.

20 Figur 8 zeigt ebenfalls ein beispielhaftes Listing von Firmen (hier  
z.B. Anwaltskanzleiseiten etc.) in Abhängigkeit von einer Selektion von  
Wertungs- und/oder Suchbegriffen 310,311,312,313 (hier z.B.  
pharmazeutische Produkte) und ihrer Trefferzahl in den Dokumenten.

Figur 9 zeigt der zeitliche Ablauf eines Ereignisses, das zu einer  
25 Class Action gegen eine Firma führen kann. Die Spezifizierung des Systems  
entsprechend diesem Ablauf ermöglicht so z.B. eine zeitliche Überwachung  
und Warnung des Benutzers vor einer möglichen und/oder wahrscheinlichen  
Class Action.

Figur 10 zeigt das Listing von Firmennamen in Abhängigkeit von Wertungsbegriffen wie z.B. Klage etc. und ihrer Trefferzahl in Nachrichten oder Emails eines Forums.

Figur 11 zeigt das Listing in der gleichen Ausführungsvariante wie in  
5 Figur 10 allgemein nach Firmennamen.

Figur 12 zeigt das Listing in der gleichen Ausführungsvariante wie in Figur 10 und 11 nach Wertungsbegriffen wie z.B. pharmazeutische Produkte.

Figur 13 zeigt ein Listing der zeitlichen Schwankung der mittel des Systems durchgeführten Aggregation und/oder Analyse der Dokumente.

10 Figur 1 illustrieren schematisch eine Architektur, die zur Realisierung der Erfindung verwendet werden kann. In diesem Ausführungsbeispiel werden zur Aggregation und Analyse von dezentralisiert gespeicherten  
Multimediadaten in einem Datenspeicher 31 ein oder mehrere verknüpfbare  
Suchbegriffe 310,311,312,313 abgespeichert. Unter Multimediadaten sind u.a.  
15 digitale Daten wie Texte, Graphiken, Bilder, Karten, Animationen, bewegte Bilder, Video, Quicktime, Tonaufnahmen, Programme (Software),  
programmbegleitende Daten und Hyperlinks oder Verweise auf Multimediadaten zu verstehen. Dazu gehören z.B. auch MPx (MP3) oder MPEGx (MPEG4 oder 7) Standards, wie sie durch die Moving Picture Experts Group definiert  
20 werden. Insbesondere können die Multimediadaten Daten im HTML- (Hyper Text Markup Language), HDML- (Handheld Device Markup Language), WMD- (Wireless Markup Language), VRML- (Virtual Reality Modeling Language) oder XML- (Extensible Markup Language) Format umfassen. Eine Recheneinheit 10 greift über ein Netzwerk 50 auf mit Quelldatenbanken 401,411,421,431  
25 verbundene Netzwerknodes 40,41,42,43 zu und Daten der Quelldatenbanken 401,411,421,431 werden basierend auf den Suchbegriffen 310,311,312,313 selektiert. Gemäss der vorliegenden Erfindung ist die Recheneinheit 10 mit den Netzwerknodes 40,41,42,43 über ein Kommunikationsnetz bidirektional verbunden. Das Kommunikationsnetz 50 umfasst beispielsweise ein GSM- oder  
30 ein UMTS-Netz, oder ein satellitenbasiertes Mobilfunknetz, und/oder ein oder mehrere Festnetze, beispielsweise das öffentlich geschaltete Telefonnetz, das



(310,311,312,313) und/oder einer Verknüpfung von Suchbegriffen (310,311,312,313) zuordbar ist,

5 dass die Recheneinheit (10) ein Filtermodul (30) zum Erzeugen einer Wertungsliste (330,331,332) von in den Quelldatenbanken (401,411,421,431) der Netzwerknodes (40,41,42,43) gefundenen Datensätzen umfasst, und

10 dass die Recheneinheit (10) ein Parametrisierungsmodul (20) zum mindestens teilweise dynamischen Generieren einer variable Stimmungsgrösse (21) basierend auf der Wertungsliste (330,331,332) für den jeweiligen Wertungsparameter (320,321,322) umfasst, welche variable Stimmungsgrösse (21) positiven und/oder negativen Stimmungsschwankungen von Benutzern des Netzwerkes (50) entsprechen.

13. System nach Anspruch 12, dadurch gekennzeichnet, dass die Recheneinheit (10) eine lexikographische Wertungsdatenbank zur Generierung eines oder mehrerer der Wertungsparameter (320,321,322) umfasst.

15 14. System nach einem der Ansprüche 12 oder 13, dadurch gekennzeichnet, dass die Recheneinheit ein Modul zur dynamischen Generierung von einem oder mehreren der Wertungsparameter (320,321,322) während der Erzeugung der Wertungsliste (330,331,332) umfasst.

20 15. System nach einem der Ansprüche 12 bis 14, dadurch gekennzeichnet, dass die Wertungsliste (330,331,332) mit den gefundenen Datensätzen und/oder Verweisen auf gefundene Datensätze in einem Contentmodul (60) der Recheneinheit (10) für einen Benutzer zugreifbar abgespeichert ist.

25 16. System nach Anspruch 15, dadurch gekennzeichnet, dass mittels der Recheneinheit (10) periodisch die Stimmungsgrössen (21) überprüfbar sind und, falls mindestens eine der Stimmungsgrössen (21) ausserhalb einer festlegbaren Schwankungstoleranz oder bestimmbaren Erwartungswertes liegt, die entsprechende Wertungsliste (330,331,332) mit